DOCUMENT RESUME

ED 115 570 SO 008 802

AUTHOR Newton, Richard F.
TITLE A Discrepancy Based Evaluation Model for International Education.
PUB DATE Nov 75
NOTE 20p.; Paper presented at the Annual Meeting of the National Council for the Social Studies (Atlanta, Georgia, November 27-29, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS *Attitudes; Change Strategies; Curriculum Development; Evaluation Methods; Foreign Culture; *Goal Orientation; International Education; *International Programs; Measurement Techniques; *Models; Need Gratification; Non Western Civilization; Personality Assessment; *Program Evaluation; Social Science Research; Test Construction
IDENTIFIERS Egypt

ABSTRACT

A discrepancy-based evaluation model for international education, developed for a program revolving around a seminar held in Egypt in 1974, is presented. Based on Malcom M. Provus' model, it considered the special problems fo a six-week educational seminar in a non-Western culture, and it questioned what type of individuals can best develop and utilize curriculum resources from another culture. The major problem was that the participants' attitudes and goals conflicted with the project's goals. A questionnaire was formulated and administered several times during the program; a free-response instrument was felt best suited for the evaluation model. It measured attitude by asking for completion of sentences such as "Egyptian people are..." and ranking responses as positive, neutral, or negative. Also, participant profiles were acquired that measured locus of control, needs achievement, and cognitive development. Results showed that the evaluation model did not answer the question for which it was designed, possibly because the model was not correctly developed for projecting that information or the question was inappropriate. It was concluded that an international program should be altered to fit participants rather than selecting "ideal" participants for a predesigned program. (ND)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED115570 .1.

SP 008 802

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A DISCREPANCY BASED EVALUATION MODEL FOR INTERNATIONAL EDUCATION

Richard F. Newton
Temple University

Presented at the College and University Faculty Assembly of the National Council for the Social Studies, November 1975

A DISCREPANCY BASED EVALUATION MODEL FOR INTERNATIONAL EDUCATION*

One of the popular activities for teachers during the summer is an educational program in foreign culture. USOE funds a number of curriculum development seminars under its Group Projects Abroad program, and a host of universities offer similar programs around the world. Considering the number of these programs it is unbelievable that no one has really developed an adequate evaluation model for them. As a result of this failure no one really knows if these activities do anything worthwhile, or even more astonishing, no one really knows if they do anything at all.

This situation confronted the author in planning a seminar to be held in the Arab Republic of Egypt. This seminar was jointly sponsored by Temple University and the Bucks County Intermediate Schools Unit. There was no evaluation model that had been used in such previous endeavors nor would any program evaluation model from other fields readily lend itself for use. The first task therefore, was to define what it was that an evaluation model should do. Obviously it should give some indication as whether or not the program met its stated purpose. At the same time there was the question as to whether or not the program met with the satisfaction of its participants. While these are two different questions, one is obviously prior to the other. The chances of a program's having met its objectives while the participants were dissatisfied were

*The author wishes to thank Dr. David Kapel of Temple(now at the University of Nebraska at Omaha) for his help in developing and refining the model used in this paper. His efforts in forcing the author to clarify his thoughts were invaluable. Likewise, Mr. H. Richard Knipple of Bucks County Intermediate Schools Unit #22 was invaluable in organizing and developing the seminar.

rather remote.

This meant that any evaluation program had to tap the expectations of the participants and incorporate these into the overall project goals. In effect the need was to take the goals of a curriculum development project, the individual expectations of twenty-four teachers, and develop a seminar program that was consistent with the goals of all concerned. The fact that the experience was to take place in a non-Western culture added to the confusion. The complexity of the situation in an international experiences seminar demanded an evaluation model that could be used for formative purposes.

The final goal of the seminar was to produce curriculum materials for students and teachers to use. The problems of dealing with curricular evaluation were rather traditional. Any evaluation program that called for assessing student outcomes would be appropriate. While appreciating the problems of curriculum evaluation, it is fair to say that it is an easier task than is program evaluation.

One of the more recent ideas in evaluation that appeared promising as work was begun on the project was the discrepency evaluation model as developed by Malcom M. Provus.[1] The most interesting aspect of this model was that it provided for both program development and program stabilization. Another attractive feature was that it assumed that programs are often times installed with inadequate consideration

---

[1]Malcom M. Provus, "The Discrepency Evaluation Model," in Readings in Curriculum Evaluation. Taylor and Cowley, eds. Dubuque, Iowa: Wm. C. Brown Co. Publishers, 1972.

of constraints and limitations. This point is important. Even the most meticulously planned program has a tendency to suffer when run in a foreign culture. A program designed in this country faces real problems in a non-Western culture. No matter how magnificant the proposal and the pre-planning has been, there will be breakdowns when the program is actually implemented. Here was where a discrepency based evaluation model appeared most helpful as it assumed the program would encounter problems and changes.

Another strong feature of this model is the assumption of differing goals by the different parties involved. Usually there are at least three types of goals involved in any such project. First, the project directors have something in mind. Next, certain goals are stated in the proposal for funding. These may be quite different from the one's the project management have, but most funded projects pay great attention (and lip-service) to the funding agency's goals whether they believe in them or not. Thirdly, the participants all have personal sets of goals. While the three sets of goals (participants', project managements', and the funding agency's) are generally similar in respects, they can be different enough to cause problems in program operation. An evaluation model which does not address itself to this problem of goal discrepencies will become part of the problem it is supposed to remedy.

In this particular project there was agreement between the goals of the project's directors and the funding proposal, but

relatively little agreement among the participants. While the avowed purpose of the program was the development of curriculum materials none of the teacher participants stated that as an individual goal when asked to state their personal objectives. Rather, they stated a predominance of what might best be called personal goals.

These included such things as making new friends, finding out what it is like to be a foreigner, enjoying the companionship of a group, gaining new insights into one's own character, and even learning Arabic. While one could probably make a case for each one of these personal objectives being related to the problem of curriculum development, they obviously did not contribute to the project's goals directly. The problem for the program directors was to determine how to incorporate these personal goals into the program design.

One of the problems in international education is the difference in resources between the United States and other nations. This has a tremendous effect on program design. The resources of a foreign culture can not be controlled in the same way that a researcher controls his dependent variables.

One way in which the information on personal objectives can be used is for consensus building. It was not done in this particular case but I would suggest that it be done in future cases. If the project participants, in the course of discussion, could come to some agreement as to the purposes of the program, the program would be helped tremendously. Any strong opinions can be analyzed within the

group and the purposes of the program can be greatly clarified. This also builds a commitment to the program. A commitment will become very important as the novelty of a new culture wears thin and exasperation grows over minor cultural differences.

It is in this first stage that a sound basis for program refinement and consistency is built. Too often in international programs the problems of living in another culture overwhelm the program, and mere continued existence becomes a primary goal. The purposes of the program are lost and people run about randomly soaking up or avoiding the culture. This is especially so when one is dealing with a non-Wester culture.

The second stage in this modified discrepency model compares actual program performance with the program outlined in Stage 1. Here the major task is one of monitoring both how the total program as well as the individual participants are preceeding with the tasks assigned them. In a program such as the Egyptian one, this meant that project directors had to be constantly aware of what resources each participant needed, and then communicate these requests to their Egyptian counterparts. Usually this meant arranging for meetings with government officials, finding out how specific experiences such as home visits could be made available for the participants, and arranging for follow-up with certain speakers.

The development of a Participant Profile was a second question which was of interest to the project directors; i.e., what types of

of people participate in an overseas seminar. This was not a lofty research question but one with a great deal of practical significance. The basic concern was: What sorts of people can benefit most from such an experience? Are there certain people who can utilize resources found in a foreign culture for developing curriculum materials more readily than others?

In asking these questions and developing an evaluation program that would allow some answers, the project utilized more of an action research style than an experimental, or even quasi-experimental, approach. The desire was simply to learn something about the participants and how they utilized resources for curriculum development. No concern was given to tight statistical designs and exhaustive psychological analyses. While this could be done there was, a real question regarding how productive such a course of action might be.

Thus, there were two factors that called for measurement in this program. The first was the problem of program evaluation. The second called for the construction of a participant profile. In selecting measurement instruments for the program, consideration was given to those instruments which could be used in both areas. The concern was one of participant fatigue. All of the pre and post measurements must be used at rather specific times in the project and teachers get so tired of taking tests very quickly. In fact this group thought the testing was a little overdone and quickly voiced their concern. As is obvious in the next section part of the problem was the uncertainty of what assessment instruments would provide the desired information.

Instrumentation of the Model

Expectations:

Of all the measures used in this program this was probably the most significant. It was the one item that was constantly reviewed during the actual seminar in Egypt. All that was asked was for each participant to list the three things he hoped to accomplish as a result of the seminar in Egypt. The question was simply stated "As a result of this experience, I expect that I will accomplish the following things:"

Some the individual expectations are listed below as examples.

Make friends -- People to People

Improve my mind set regarding the Middle East

Gain a sense of appreciation for a people of the non-Christian faith

A positive feeling towards the Egyptian people and the Muslim faith

I hope to make new friends

The most interesting thing about these responses was that they all stated a personal goal. Out of roughly a hundred different stated expectations only three or four mentioned students. While one can easily argue that by changing a teacher's knowledge and attitudes, students will in turn be effected, it is nonetheless interesting that so few mentioned teaching and students in their expectations.

As problems arose in the course of the seminar, or as displeasure was expressed with certain features of the seminar, the expectation sheets were used the way that a counselor might use a student auto-

biography. Likewise, the expectation sheets were treated as private documents, and the participants did not share them with each other.

While these statements were extremely useful in the 1974 seminar they can prove to be even more valuable in future seminars. One way would be to ask the participants to share their expectations with each other. This in no way attempts to force people to adhere to a common set of expectations through the use of peer pressure. Rather it should be used as a way of considering what others hope to accomplish by way of hearing their rationale. In using a discrepency based evaluation model this is very important.

Attitude Measures

The most important point to keep in mind when discussing the affective area is that one is not really expecting a change of attitudes. The reason for this is that prior to visiting another culture it would be unusual to have anything less than a positive attitude towards the culture. At least this is generally true of teachers who participate in group project abroad programs. It might not be true of individuals who work for large multinational corporations or people who work for government agencies. Often times these people have no choice as to where they go in the world. Teachers do have this choice. They make a conscious decision to spend a summer living in a different culture. Hence they already have favorable inclination towards that culture; at the very least they might be neutral. It would be rare to find a teacher actually hostile towards

that culture.

With all of this in mind an effort was made to develop an instrument which could ass ss a change in magnitude rather than a change in direction. A free-response instrument was thought to be best suited, and a sentence completion format was utilized. The participants were asked to complete the following sentences.

1. Egyptian people are _____
2. Egyptian music is _____
3. Egyptian industry is _____
4. Egyptian education is _____
5. The Egyptian government is _____
6. Modernization in Egypt is _____
7. Women in Egypt are _____
8. The arts in Egypt are _____
9. Villages in Egypt are _____
10. Religion in Egypt is _____
11. Social Services in Egypt are _____
12. Egyptian's attitude toward Americans is _____

The response(s) the teacher gave was examined for the type of descriptor used. These descriptors were classified as being either negative, positive, or neutral. An example of a descriptor is "Egyptian Government is not democratic" or "Social services in Egypt are not very well established." In both cases these would be classified as negative descriptors. A phase such as "The Arts in Egypt are exceptional and

and unique," would be classified as positive as would "Social services in Egypt are advanced." Phrases such as "Egyptian music is based on a pattern of repetition," and "Egyptian industry is varied" were scored as neutral responses.

On the pre-test 67 responses were classified as negative, 66 as positive and 50 as neutral. In the post-test situation 65 responses were classified as negative, 268 as positive and 24 as neutral. Thus, there is a dramatic increase in the number of positive phrases being used to complete the sentences, an increase of over four times as many positive descriptors. Also, only one-half as many neutral responses were found. This would indicate that whatever suspended judgments the participants had regarding Egypt prior to the seminar were changed to positive ones as a result of the experience.

An interesting result is that the number of negative responses did not decrease. The explanation here may well be that people willing to spend six weeks in Egypt did not hold a negative attitude towards the country. If anything, they were as positive towards Egypt as possible. In the process of the seminar whatever neutral attitudes the teachers held were converted to positive ones. In addition, a tremendous amount of new positive attitudes were developed.

Examples of some of these positive attitudes are listed below:

Egyptian people are:

sincere, warm and open
warm, hospitable and very proud
friendly, hard-working, religious human beings
friendly, sincere, hard-working, and proud of their country.

Participant Profile

Another important aspect of the evaluation program as carried out in this project, was an attempt to create the previously mentioned participant profile. All of the tests described below were used to create just such a profile. Before describing the instruments it is important to point out that the project was never intended to alter behavior. Anyone who would argue that a six-week program in Egypt will permanently alter a person's behavior in such critical areas as locus of control, needs achievement, and dogmatism, ought to confine his activities to writing books for the best seller list. All of the following tests were designed to give some indication as to what the participants in such a program look like psychologically.

Locus of Control:

One of the classic two category systems for separating people is the internal control-external control diochotomy. While it is foolish to assume that a person either fits into one of these categories or the other, it is not foolish to assume that people do differ, to some degree, in which they are directed by the expectations of others. The internally controlled person is thought of as being highly self-sufficient and independent while the person with a high degree of external control is the individual who always waits for his cues from others.

This variable is an important one to know about the seminar participants. We began with the idea that a person who ranks high in either direction will present problems in a group seminar project. The person who is highly internally motivated will be a problem in

that he cannot subordinate his personal desires and needs. He is just too independent. On the other hand the person who is highly externally motivated will present problems as he will always be waiting for others to tell him what is appropriate. Essentially he will not move unless someone tells him to. One of the hypotheses that was entertained prior to the seminar was that participants on either end of the scale would present problems. The reason for this was that the seminar was essentailly a group activitiy and that anyone who was highly inner directed would not long stand for the group process of doing things. Likewise a person who was highly externally directed would present a problem in that he would not do anything unless he was told to do so. The ideal participant, it was envisioned, would be a person one standard deviation from the mean. Indeed this proved to be the case. When the overall effectiveness of the participants was assessed, those who fell lowest were those people who were on the extremes of the laws of control scale. One participant scored as high as is possible on inner-directedness. As a member of the group he worked very hard to participate. At certain intervals though he would disappear from the group and pursue his own interests. Also, once he returned to the United States he had great difficulty in working as a group member in the development of curriculum materials.

Another participant was scored equally high towards the other directed end of the scale. This person did present more of a problem. He fell into the control of another member of the group and together they filed

minority reports on every aspect of the seminar. Also this person had great difficulty in working on his curriculum materials in Egypt and kept wanting to change his topic to match what someone else was doing.

With the seminar organized the way it was - a group endeavor - those people who represented a blend of inner and external control functioned best. In the summary section some suggestions are given for dealing with participants in other ways, but the conclusion nonetheless holds that if the program is organized on a group basis then participants who have strong tendencies towards either high external or internal control experience difficulty.

Needs Achievement:

One of the variables that it was assumed would make a difference was the psychological need that people had to be successful. The instrument that was used allowed the overall score to be broken into two component parts: need to achieve, and need to avoid failure. This is a basic conceptual difference with regards to the average needs achievement test.

Oftentimes it appears that the traditional conception of achievement is simply too narrow. Many people do not work hard to be a success, but rather work hard to avoid failure, yet it appears they are success oriented. It is similar to the former Green Bay Packer who pointed out that everybody wants to win, the difference is that championship players just hate to lose worse. Although all of these considerations were made, it was found that the needs achievement orientation of an individual had no correlation with his

overall effectiveness. This was not because of homogeneity either. The participants ranged from an almost total lack of need for achievement to those who could best be classified as compulsive achievers (need to achieve, need to avoid failure). Just as with the Locus of Control measure, the group was normally distributed. The only conclusion was that the need for achievement neither detracts from nor contributed to the effectiveness of a participant in a seminar such as the Egyptian one held in 1974.

Cognitive Tests:

The interest in the cognitive domain is rather obvious. A person who spends six weeks in a foreign culture should demonstrate a gain in his knowledge of that culture. This is the least that can be expected from that experience. It is also the most easily measured of all the items of possible interest.

In the 1974 seminar a simple true-false test was used. The only reason for the use of a true-false format was that it was easier to achieve reliability with a true-false test than it was with a multiple choice test. Test reliability is a factor of test length and true-false test items are easier to write than are multiple choice items.

The experience of this seminar demonstrated that people do indeed gain substantially in knowledge of a country as a result of residing in it. The question needs to be asked though if the gain is such as to justify the high cost of an international experience. Could a person learn just as much information by reading an appropriate number of books? The answer is probably yes.

At any rate this is not the point. Indeed, one does learn a great deal of information from visiting a foreign culture -- this no one doubts. Were there no significant increase in cognitive gain scores, one should surprised; and the viability of the program should be seriously questioned. A program such as the one being described here can hardly be justified on the basis of a gain in cognitive knowledge but this gain is nonetheless important. A comparison of the international seminars cost-effectiveness with alternative programs is a separate point and one that is not examined in this paper.

Materials Evaluation

A feature of the evaluation program that has not been discussed is the assessment of the curriculum materials produced by the participants. This phase of the evaluation program is rather conventional in that it attempts to measure the change in student knowledge resulting from a treatment. The treatment consisted of exposure to a specific set of learning materials developed by the seminar participants.

The question of significance here is whether the evaluation should be based on a comparative or non-comparative basis. Many in curriculum would argue that curricula should be evaluated on the ends that it seeks to attain. The process is simply one of determining which of the specified curriculum objectives were attained. Any attempt to compare curriculum A with curriculum B is unfortunate, and misguided, since the two curricula obviously do not have identical objectives. In

addition curricula are seldom static. They are constantly evolving and changing. An evaluation only gives us a picture of where things are at a given point in time. Any revisions are likely to change the ranking.

The only problem with this is that the argument is aimed at the curriculum developers not the curriculum selector. In this particular project the interest is in whether the materials developed by teachers who spent a summer in Egypt are more realistic and valid than the commercial alternatives. Comparison on this basis is entirely legitimate.

The materials developed as a result of this seminar are presently being evaluated on a non-comparative evaluation such a comparison is beyond the scope of the project.

Summary and Recommendations

This evaluation model raises two central points. The first is that the model did not answer the question: "What type of individuals could best develop and utilize curriculum resources which reside in a different culture?" First, the evaluation model was not well developed enough to project that kind of information. Secondly, as the total evaluation model is examined, there is presently some doubt that the question was appropriate. What began as an attempt to select participants for a pre-designed program has been revised to the point to where a design for program change is probably more appropriate.

The hope at the start of the project was that the participant profile could be used in selecting people for future seminars. That

is not viewed as a naive idea. The task of predicting who can, and who cannot, benefit from the Group Projects Abroad program is far too complex a question. Outside of a few basic criteria such as physical health and stamina no flat guidelines can be laid down.

The more productive course of action is to use the participant profiles for making changes in program design. If, for instance, one is dealing with a group of educators who are highly internally controlled then a great deal of thought should be given to the individualization of as much of the program as is possible. Likewise a group of people who have a set of unrealistic expectations ought to be placed in a situation where both they, and the others involved in the project, can deal with these expectations.

Basically what began as an attempt to manipulate the participant selection process ended with the conclusion that the program variables ought to be manipulated. This also increases the insurance that the program <u>will</u> be a success. Indeed, once a decision is made to utilize a discrepency evaluation model, the program should be a guaranteed success.

The reason for this is that the program is constantly manipulated so as to insure that it meets the goals of the parties involved. This is <u>not</u> the same as manipulating the data in an effort to demonstrate success. Data manipulation is obviously fraudulent. Program manipulation is only common sense. In a discrepency based evaluation model it is done in a systematic manner.

Another important point to be kept in mind when using discrepency based evaluation is that instrumentation is extremely important. It is most important that the data include information that will allow for formative evaluation. Program change is a central feature of the model. Any instruments that are utilized must be ones that contribute to making sound decisions about such program changes.

In this area the most important data that can be collected are the expectations that the different participants bring to the program. Variations between these individual goals or expectations need to be dealt with through group consensus and goal clarification. It is important to keep in mind that this will not involve major shifts, i.e., a curriculum development seminar will not become an encounter group, but rather it will involve minor programmatic adjustments. These adjustments are crucial to the functioning of a successful international experience.